

RESEARCH

Open Access

Characterizing compatibility and agreement of unrooted trees via cuts in graphs

Sudheer Vakati and David Fernández-Baca*

Abstract

Background: Deciding whether there is a single tree—a supertree—that summarizes the evolutionary information in a collection of unrooted trees is a fundamental problem in phylogenetics. We consider two versions of this question: agreement and compatibility. In the first, the supertree is required to reflect precisely the relationships among the species exhibited by the input trees. In the second, the supertree can be more refined than the input trees. Testing for compatibility is an NP-complete problem; however, the problem is solvable in polynomial time when the number of input trees is fixed. Testing for agreement is also NP-complete, but it is not known whether it is fixed-parameter tractable. Compatibility can be characterized in terms of the existence of a specific kind of triangulation in a structure known as the display graph. Alternatively, it can be characterized as a chordal graph sandwich problem in a structure known as the edge label intersection graph. No characterization of agreement was known.

Results: We present a simple and natural characterization of compatibility in terms of minimal cuts in the display graph, which is closely related to compatibility of splits. We then derive a characterization for agreement.

Conclusions: Explicit characterizations of tree compatibility and agreement are essential to finding practical algorithms for these problems. The simplicity of the characterizations presented here could help to achieve this goal.

Keywords: Phylogenies, Supertrees, Compatibility, Agreement, Cuts in graphs, Chordal graphs

Background

A *phylogenetic tree* T is an unrooted tree whose leaves are bijectively mapped to a label set $\mathcal{L}(T)$. Labels represent species and T represents the evolutionary history of these species. Let \mathcal{P} be a collection of phylogenetic trees. We call \mathcal{P} a *profile*, refer to the trees in \mathcal{P} as *input trees*, and denote the combined label set of the input trees, $\bigcup_{T \in \mathcal{P}} \mathcal{L}(T)$, by $\mathcal{L}(\mathcal{P})$. A *supertree* of \mathcal{P} is a phylogenetic tree whose label set is $\mathcal{L}(\mathcal{P})$. The goal of constructing a supertree for a profile is to synthesize the information in the input trees in a larger, more comprehensive, phylogeny [1]. Ideally, a supertree should faithfully reflect the relationships among the species implied by the input trees. In reality, this is rarely achievable, because of conflicts among the input trees due to errors in constructing them or to biological processes such as lateral gene transfer and gene duplication.

We consider two classic versions of the supertree problem, based on the closely related notions of compatibility

and agreement. Let S and T be two phylogenetic trees where $\mathcal{L}(T) \subseteq \mathcal{L}(S)$ —for our purposes, T would be an input tree and S a supertree. Let S' be the tree obtained by suppressing any degree-two vertices in the minimal subtree of S connecting the labels in $\mathcal{L}(T)$. We say that S *displays* T , or that T and S are *compatible*, if T can be derived from S' by contracting edges. We say that tree T is an *induced subtree* of S , or that T and S *agree*, if S' is isomorphic to T .

Let \mathcal{P} be a profile. The *tree compatibility problem* asks if there exists a supertree for \mathcal{P} that displays all the trees in \mathcal{P} . If such a supertree S exists, we say that \mathcal{P} is *compatible* and S is a *compatible supertree* for \mathcal{P} . The *agreement supertree problem* asks if there exists a supertree for \mathcal{P} that agrees with all the trees in \mathcal{P} . If such a supertree S exists, we say that S is an *agreement supertree* (AST) for \mathcal{P} .

Compatibility and agreement embody different philosophies about conflict. An agreement supertree must reflect precisely the evolutionary relationships exhibited by the input trees. In contrast, a compatible supertree is allowed to exhibit more fine-grained relationships among certain

*Correspondence: fernande@iastate.edu

Department of Computer Science, Iowa State University, Ames, Iowa, USA

labels than those exhibited by an input tree. From a biological viewpoint, the differences between compatibility and agreement reflect different ways to treat *polytomies* —i.e., nodes of degree greater than three. Compatibility treats polytomies as *soft* facts: if an input tree node has degree four or more, it is not because there were multiple simultaneous speciation events, but because there is not enough information to resolve the sequence of speciation. Thus, if another input tree provides more refined information about speciation order, we can use it, provided the information is not contradicted by the remaining input trees. Agreement, in contrast, treats polytomies as *hard* facts. Note that compatibility and agreement are equivalent when the input trees are binary.

If all the input trees share a common label (which can be viewed as a root node), both tree compatibility and agreement are solvable in polynomial time [2,3]. In general, however, the two problems are NP-complete, and remain so even when the trees are quartets; i.e., binary trees with exactly four leaves [4]. Nevertheless, Bryant and Lagergren showed that the tree compatibility problem is fixed-parameter tractable when parametrized by number of trees [5]. It is unknown whether or not the agreement supertree problem has the same property.

To prove the fixed-parameter tractability of tree compatibility, Bryant and Lagergren first showed that a necessary (but not sufficient) condition for a profile to be compatible is that the tree-width of a certain graph —the *display graph* of the profile (see Section ‘Display graphs and edge label intersection graphs’)— be bounded by the number of trees. They then showed how to express compatibility as a bounded-size monadic second-order formula on the display graph. By Courcelle’s Theorem [6,7], these two facts imply that compatibility can be decided in time linear in the size of the display graph. Unfortunately, Bryant and Lagergren’s argument amounts essentially to only an existential proof, as it is not clear how to obtain an explicit algorithm for unrooted compatibility from it.

A necessary step towards finding a practical algorithm for compatibility —and indeed for agreement— is to develop an explicit characterization of the problem. In earlier work [8], we made some progress in this direction, characterizing tree compatibility in terms of the existence of a legal triangulation of the display graph of the profile. Gysel et al. [9] provided an alternative characterization, based on a structure they call the edge label intersection graph (ELIG) (see Section ‘Display graphs and edge label intersection graphs’). Their formulation is in some ways simpler than that of [8], allowing Gysel et al. to express tree compatibility as a chordal sandwich problem. Neither [8] nor [9] deal with agreement.

Here, we show that the connection between separators in the ELIG and cuts in the display graph (explored

in Section ‘Display graphs and edge label intersection graphs’) leads to a new, and natural, characterization of compatibility in terms of minimal cuts in the display graph (Section ‘Characterizing compatibility via cuts’). We then show how such cuts are closely related to the splits of the compatible supertree (Section ‘Splits and cuts’). Next, we give a characterization of the agreement in terms of minimal cuts of the display graph (Section ‘Characterizing agreement via cuts’). To our knowledge, there was no previous characterization of the agreement supertree problem for unrooted trees. Lastly, we examine the connection between the triangulation-based and the cut-based perspectives on compatibility (Section ‘Relationship to legal triangulations’).

Preliminaries

Splits, compatibility, and agreement

A *split* of a label set L is a bipartition of L consisting of non-empty sets. We denote a split $\{X, Y\}$ by $X|Y$. A split is *non-trivial* if neither of its sets is a singleton; otherwise, it is *trivial*. Let T be a phylogenetic tree. Let e be an edge of T . Deletion of e disconnects T into two subtrees T_1 and T_2 . If L_1 and L_2 denote the set of all labels in T_1 and T_2 , respectively, then $L_1|L_2$ is a split of $\mathcal{L}(T)$. We denote by $\sigma_e(T)$ the split corresponding to edge e of T ; if e is a leaf edge, then $\sigma_e(T)$ is a trivial split. Let $\Sigma(T)$ denote the set of all splits corresponding to internal edges of T and $\Sigma_{\text{triv}}(T)$ denote the set of all (trivial) splits corresponding to leaf edges of T .

A tree T *displays* a split $X|Y$ if there exists an internal edge e of T where $\sigma_e(T) = X|Y$. A set of splits of a label set L is *compatible* if there exists a tree that displays all the splits in the set. It is well-known that two splits $A_1|A_2$ and $B_1|B_2$ are compatible if and only if at least one of $A_1 \cap B_1$, $A_1 \cap B_2$, $A_2 \cap B_1$ and $A_2 \cap B_2$ is empty [10]. Note that a trivial split of L is compatible with every split of L .

Theorem 1 (Splits-Equivalence Theorem [10,11]). *Let Σ be a collection of splits of a label set X that includes all trivial splits. Then, $\Sigma = \Sigma(T) \cup \Sigma_{\text{triv}}(T)$ for some phylogenetic tree T with label set X if and only if the splits in Σ are pairwise compatible. Tree T is unique up to isomorphism.*

Let S be a phylogenetic tree and let Y be a subset of $\mathcal{L}(S)$. Then, $S|_Y$ denotes the tree obtained by suppressing any degree-two vertices in the minimal subtree of S connecting the labels in Y . Now, let T be a phylogenetic tree such that $\mathcal{L}(T) \subseteq \mathcal{L}(S)$. Then, S *displays* T if and only if $\Sigma(T) \subseteq \Sigma(S|_{\mathcal{L}(T)})$; T and S *agree* if and only if $\Sigma(T) = \Sigma(S|_{\mathcal{L}(T)})$.

Cliques, separators, cuts, and triangulations

Let G be a graph. We represent the vertices and edges of G by $V(G)$ and $E(G)$ respectively. A *clique* of G is a complete

subgraph of G . A clique H of G is *maximal* if there is no other clique H' of G where $V(H) \subset V(H')$. For any $U \subseteq V(G)$, $G-U$ is the graph derived by removing vertices of U and their incident edges from G . For any $F \subseteq E(G)$, $G-F$ is the graph with vertex set $V(G)$ and edge set $E(G) \setminus F$.

For any two nonadjacent vertices a and b of G , an a - b separator of G is a set U of vertices where $U \subset V(G)$ and a and b are in different connected components of $G-U$. An a - b separator U is *minimal* if for every $U' \subset U$, U' is not an a - b separator. A set $U \subseteq V(G)$ is a *minimal separator* if U is a minimal a - b separator for some non-adjacent vertices a and b of G . We represent the set of all minimal separators of graph G by Δ_G . Two minimal separators U and U' are *parallel* if $G-U$ contains at most one component H where $V(H) \cap U' \neq \emptyset$.

A connected component H of $G-U$ is *full* if for every $u \in U$ there exists some vertex $v \in H$ where $\{u, v\} \in E(G)$.

Lemma 1 ([12]). *For a graph G and any $U \subset V(G)$, U is a minimal separator of G if and only if $G-U$ has at least two full components.*

A *chord* is an edge between two nonadjacent vertices of a cycle. A graph H is *chordal* if and only if every cycle of length four or greater in H has a chord. A chordal graph H is a *triangulation* of graph G if $V(G) = V(H)$ and $E(G) \subseteq E(H)$. The edges in $E(H) \setminus E(G)$ are called *fill-in* edges of G . A triangulation is *minimal* if removing any fill-in edge yields a non-chordal graph.

A *clique tree* of a chordal graph H is a pair (T, B) where (i) T is a tree, (ii) B is a bijective function from vertices of T to maximal cliques of H , and (iii) for every vertex $v \in H$, the set of all vertices x of T where $v \in B(x)$ induces a subtree in T . Property (iii) is called *coherence*.

Let \mathcal{F} be a collection of subsets of $V(G)$. We represent by $G_{\mathcal{F}}$ the graph derived from G by making the set of vertices of X a clique for every $X \in \mathcal{F}$. The next result summarizes basic facts about separators and triangulations (see [12-14]).

Theorem 2. *Let \mathcal{F} be a maximal set of pairwise parallel minimal separators of G and H be a minimal triangulation of G . Then, the following statements hold.*

1. $G_{\mathcal{F}}$ is a minimal triangulation of G .
2. Let (T, B) be a clique tree of $G_{\mathcal{F}}$. There exists a minimal separator $F \in \mathcal{F}$ if and only if there exist two adjacent vertices x and y in T where $B(x) \cap B(y) = F$.
3. Δ_H is a maximal set of pairwise parallel minimal separators of G and $G_{\Delta_H} = H$.

A *cut* in a connected graph G is a subset F of edges of G such that $G-F$ is disconnected. A cut F is *minimal* if there does not exist $F' \subset F$ where $G-F'$ is disconnected. Note

that if F is minimal, $G-F$ has exactly be two connected components. Two minimal cuts F and F' are *parallel* if $G-F$ has at most one connected component H where $E(H) \cap F' \neq \emptyset$.

Display graphs and edge label intersection graphs

We now introduce the two main notions that we use to characterize compatibility and agreement: the display graph and the edge label intersection graph. We then present some known results about these graphs, along with new results on the relationships between them. Here and in the rest of the paper, $[m]$ denotes the set $\{1, \dots, m\}$, where m is a positive integer. Since for any phylogenetic tree T there is a bijection between the leaves of T and $\mathcal{L}(T)$, we refer to the leaves of T by their labels.

Let $\mathcal{P} = \{T_1, T_2, \dots, T_k\}$ be a profile. We assume that for any $i, j \in [k]$ such that $i \neq j$, the sets of internal vertices of input trees T_i and T_j are disjoint. The *display graph* of \mathcal{P} , denoted by $G(\mathcal{P})$, is a graph whose vertex set is $\bigcup_{i \in [k]} V(T_i)$ and edge set is $\bigcup_{j \in [k]} E(T_j)$ (see Figure 1). A vertex v of $G(\mathcal{P})$ is a *leaf* if $v \in \mathcal{L}(\mathcal{P})$. Every other vertex of $G(\mathcal{P})$ is an *internal*. An edge of $G(\mathcal{P})$ is *internal* if its endpoints are both internal.

A triangulation G' of $G(\mathcal{P})$ is *legal* if it satisfies the following conditions.

1. For every clique C of G' , if C contains an internal edge, then it contains no other edge of $G(\mathcal{P})$.
2. No fill-in edge in G' has a leaf as an endpoint.

Theorem 3 (Vakati, Fernández-Baca [8]). *A profile \mathcal{P} of unrooted phylogenetic trees is compatible if and only if $G(\mathcal{P})$ has a legal triangulation.*

In what follows, we assume that $G(\mathcal{P})$ is connected. If it is not, the connected components of $G(\mathcal{P})$ induce a partition of \mathcal{P} into sub-profiles such that for each sub-profile \mathcal{P}' , $G(\mathcal{P}')$ is a connected component of $G(\mathcal{P})$. It is easy to see that \mathcal{P} is compatible if and only if each sub-profile is compatible.

The *edge label intersection graph* of \mathcal{P} , denoted $LG(\mathcal{P})$, is the line graph of $G(\mathcal{P})$ [9]. That is, the vertex set of $LG(\mathcal{P})$ is $E(G(\mathcal{P}))$ and two vertices of $LG(\mathcal{P})$ are adjacent if the corresponding edges in $G(\mathcal{P})$ share an endpoint. (We should note that Gysel et al. [9] refer to $LG(\mathcal{P})$ as the modified edge label intersection graph.) For an unrooted tree T , $LG(T)$ denotes $LG(\{T\})$.

Observation 1. *Let F be a set of edges of $G(\mathcal{P})$ and let $\{v_1, v_2, \dots, v_m\} \subseteq V(G(\mathcal{P}))$ where $m \geq 2$. Then, v_1, v_2, \dots, v_m is a path in $G(\mathcal{P}) - F$ if and only if $\{v_1, v_2\}, \dots, \{v_{m-1}, v_m\}$ is a path in $LG(\mathcal{P}) - F$.*

Thus, if $G(\mathcal{P})$ is connected, so is $LG(\mathcal{P})$. Hence, in what follows, we assume that $LG(\mathcal{P})$ is connected.

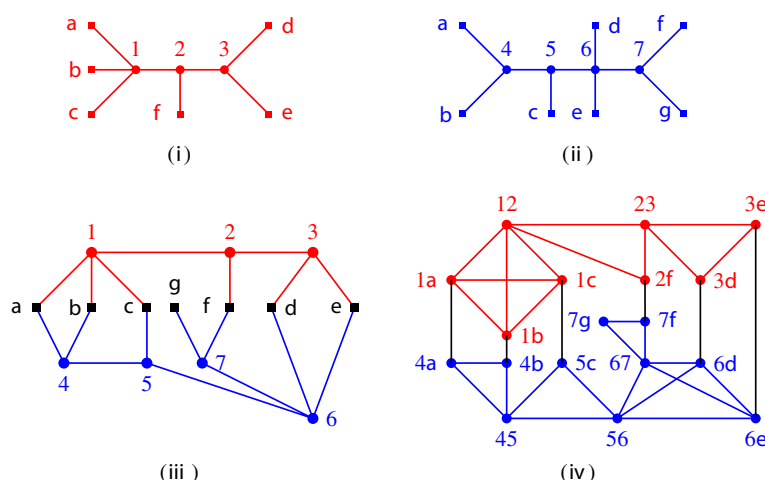


Figure 1 Compatible trees. (i) First input tree. (ii) A second input tree, compatible with the first. (iii) Display graph of the input trees. (iv) Edge label intersection graph of the input trees; for each vertex, uv represents edge $\{u, v\}$ of the display graph.

A fill-in edge for $LG(\mathcal{P})$ is *valid* if for every $T \in \mathcal{P}$, at least one of the endpoints of the edge is not in $LG(T)$. A triangulation H of $LG(\mathcal{P})$ is *restricted* if every fill-in edge of H is valid.

Theorem 4 (Gysel et al. [9]). *A profile \mathcal{P} of unrooted phylogenetic trees is compatible if and only if $LG(\mathcal{P})$ has a restricted triangulation.*

A minimal separator F of $LG(\mathcal{P})$ is *legal* if for every $T \in \mathcal{P}$, all the edges of T in F share a common endpoint; i.e., $F \cap E(T)$ is a clique in $LG(T)$. The following theorem was mentioned in [9]. For future reference, we formally state it and prove it here.

Theorem 5. *A profile \mathcal{P} is compatible if and only if there exists a maximal set \mathcal{F} of pairwise parallel minimal separators in $LG(\mathcal{P})$ where every separator in \mathcal{F} is legal.*

Proof. Our approach is similar to the one used by Gusfield in [15]. Assume that \mathcal{P} is compatible. From Theorem 4, there exists a restricted triangulation H of $LG(\mathcal{P})$. We can assume that H is minimal (if it is not, simply delete fill-in edges repeatedly from H until it is minimal). Let $\mathcal{F} = \Delta_H$. From Theorem 2, \mathcal{F} is a maximal set of pairwise parallel minimal separators of $LG(\mathcal{P})$ and $LG(\mathcal{P})_{\mathcal{F}} = H$. Suppose \mathcal{F} contains a separator F that is not legal. Let $\{e, e'\} \subseteq F$ where $\{e, e'\} \subseteq E(T)$ for some input tree T and $e \cap e' = \emptyset$. The vertices of F form a clique in H . Thus, H contains the edge $\{e, e'\}$. Since $\{e, e'\}$ is not a valid edge, H is not a restricted triangulation, a contradiction. Hence, every separator in \mathcal{F} is legal.

Let \mathcal{F} be a maximal set of pairwise parallel minimal separators of $LG(\mathcal{P})$ where every separator in \mathcal{F} is legal. From

Theorem 2, $LG(\mathcal{P})_{\mathcal{F}}$ is a minimal triangulation of $LG(\mathcal{F})$. If $\{e, e'\} \in E(LG(\mathcal{P})_{\mathcal{F}})$ is a fill-in edge, then $e \cap e' = \emptyset$ and there exists a minimal separator $F \in \mathcal{F}$ where $\{e, e'\} \subseteq F$. Since F is legal, if $\{e, e'\} \subseteq E(T)$ for some input tree T then $e \cap e' \neq \emptyset$. Thus, e and e' are not both from $LG(T)$ for any input tree T . Hence, every fill-in edge in $LG(\mathcal{P})_{\mathcal{F}}$ is valid, and $LG(\mathcal{P})_{\mathcal{F}}$ is a restricted triangulation. \square

Let u of be a vertex of some input tree, We write $\text{Inc}(u)$ to denote the set of all edges of $G(\mathcal{P})$ incident on u . Equivalently, $\text{Inc}(u)$ is the set of all vertices e of $LG(\mathcal{P})$ such that $u \in e$.

Let F be a cut of the display graph $G(\mathcal{P})$. F is *legal* if for every tree $T \in \mathcal{P}$, the edges of T in F are incident on a common vertex; i.e., if $F \cap E(T) \subseteq \text{Inc}(u)$ for some $u \in V(T)$. F is *nice* if F is legal and each connected component of $G(\mathcal{P}) - F$ has at least one edge.

Lemma 2. *Let F be a subset of $E(G(\mathcal{P}))$. Then, F is a legal minimal separator of $LG(\mathcal{P})$ if and only if F is a nice minimal cut of $G(\mathcal{P})$.*

To prove the Lemma 2, we need two auxiliary lemmas and a corollary.

Lemma 3. *Let F be any minimal separator of $LG(\mathcal{P})$ and u be any vertex of any input tree. Then, $\text{Inc}(u) \not\subseteq F$.*

Proof. Suppose F is a minimal a - b separator of $LG(\mathcal{P})$ and u is a vertex of some input tree such that $\text{Inc}(u) \subseteq F$. Consider any vertex $e \in \text{Inc}(u)$. Then, there exists a path π from a to b in $LG(\mathcal{P})$ where e is the only vertex of F in π . If such a path π did not exist, then $F - e$ would still be an a - b separator, and F would not be minimal, a contradiction.

Let e_1 and e_2 be the neighbors of e in π and let $e = \{u, v\}$. Since $\text{Inc}(u) \subseteq F$, π does not contain any other vertex e' where $u \in e'$. Thus, $e \cap e_1 = \{v\}$ and $e \cap e_2 = \{v\}$. Let $\pi = a, \dots, e_1, e, e_2, \dots, b$. Then $\pi' = a, \dots, e_1, e_2, \dots, b$ is also a path from a to b . But π' does not contain any vertex of F , contradicting the assumption that F is a separator of $\text{LG}(\mathcal{P})$. Hence, neither such a minimal separator F nor such a vertex u exist. \square

Lemma 4. *If F is a minimal separator of $\text{LG}(\mathcal{P})$, then $\text{LG}(\mathcal{P}) - F$ has exactly two connected components.*

Proof. Assume that $\text{LG}(\mathcal{P}) - F$ has more than two connected components. By Lemma 1, $\text{LG}(\mathcal{P}) - F$ has at least two full components. Let H_1 and H_2 be two full components of $\text{LG}(\mathcal{P}) - F$. Let H_3 be a connected component of $\text{LG}(\mathcal{P}) - F$ different from H_1 and H_2 . By assumption $\text{LG}(\mathcal{P})$ is connected. Thus, there exists an edge $\{e, e_3\}$ in $\text{LG}(\mathcal{P})$ where $e \in F$ and $e_3 \in H_3$. Since H_1 and H_2 are full components, there exist edges $\{e, e_1\}$ and $\{e, e_2\}$ in $\text{LG}(\mathcal{P})$ where $e_1 \in V(H_1)$ and $e_2 \in V(H_2)$.

Let $e = \{u, v\}$, and assume without loss of generality that $u \in e \cap e_3$. Then, there is no vertex $f \in V(H_1)$ where $u \in e \cap f$. Thus, $v \in e \cap e_1$. Similarly, there is no vertex $f \in V(H_2)$ such that $u \in f \cap e$ or $v \in f \cap e$. But then H_2 does not contain a vertex adjacent to e , so H_2 is not a full component, a contradiction. \square

Corollary 1. *If F is a minimal separator of $\text{LG}(\mathcal{P})$, then $\text{LG}(\mathcal{P}) - F'$ is connected for any $F' \subset F$.*

Proof of Lemma 2. We prove that if F is a legal minimal separator of $\text{LG}(\mathcal{P})$ then F is a nice minimal cut of $G(\mathcal{P})$. The proof for the other direction is similar and is omitted.

First, we show that F is a cut of $G(\mathcal{P})$. Assume the contrary. Let $\{u, v\}$ and $\{p, q\}$ be vertices in different components of $\text{LG}(\mathcal{P}) - F$. Since $G(\mathcal{P}) - F$ is connected, there is a path between vertices u and q in this graph. Also, $\{u, v\} \notin F$ and $\{p, q\} \notin F$. Thus, by Observation 1 there is also a path between vertices $\{u, v\}$ and $\{p, q\}$ of $\text{LG}(\mathcal{P}) - F$. This implies that $\{u, v\}$ and $\{p, q\}$ are in the same connected component of $\text{LG}(\mathcal{P}) - F$, a contradiction. Thus F is a cut.

Next we show that F is a nice cut of $G(\mathcal{P})$. For every $T \in \mathcal{P}$ all the vertices of $\text{LG}(T)$ in F form a clique in $\text{LG}(T)$. Thus, all the edges of T in F are incident on a common vertex, so F is legal. To complete the proof, assume that $G(\mathcal{P}) - F$ has a connected component with no edge and let u be the vertex in one such component. Then, $\text{Inc}(u) \subseteq F$. But F is a minimal separator of $\text{LG}(\mathcal{P})$, and by Lemma 3, $\text{Inc}(u) \not\subseteq F$, a contradiction. Thus, F is a nice cut.

Lastly, we show that F is a minimal cut of $G(\mathcal{P})$. Assume, on the contrary, that there exists $F' \subset F$ where $G(\mathcal{P}) -$

F' is disconnected. Since $F' \subset F$ and every connected component of $G(\mathcal{P}) - F$ has at least one edge, every connected component of $G(\mathcal{P}) - F'$ also has at least one edge. Let $\{u, v\}$ and $\{p, q\}$ be the edges in different components of $G(\mathcal{P}) - F'$. By Corollary 1, $\text{LG}(\mathcal{P}) - F'$ is connected and thus, there is a path between $\{u, v\}$ and $\{p, q\}$ in $\text{LG}(\mathcal{P}) - F'$. By Observation 1 there must also be a path between vertices u and p in $G(\mathcal{P}) - F'$. Hence, edges $\{u, v\}$ and $\{p, q\}$ are in the same connected component of $G - F'$, a contradiction. Thus, F is a minimal cut. \square

Lemma 5. *Two legal minimal separators F and F' of $\text{LG}(\mathcal{P})$ are parallel if and only if the nice minimal cuts F and F' are parallel in $G(\mathcal{P})$.*

Proof. Assume that separators F and F' of $\text{LG}(\mathcal{P})$ are parallel, but cuts F and F' of $G(\mathcal{P})$ are not. Then, there exists a set $\{\{u, v\}, \{p, q\}\} \subseteq F'$ where $\{u, v\}$ and $\{p, q\}$ are in different components of $G(\mathcal{P}) - F$. Since F and F' are parallel separators in $\text{LG}(\mathcal{P})$, and F does not contain $\{u, v\}$ and $\{p, q\}$, there exists a path between vertices $\{u, v\}$ and $\{p, q\}$ in $\text{LG}(\mathcal{P}) - F$. Then, by Observation 1 there also exists a path between vertices u and q in $G(\mathcal{P}) - F$. Thus, $\{u, v\}$ and $\{p, q\}$ are in the same connected component of $G(\mathcal{P}) - F$, a contradiction.

The other direction can be proved similarly, using Observation 1. \square

The next lemma, from [9], follows from the definition of restricted triangulation.

Lemma 6. *Let H be a restricted triangulation of $\text{LG}(\mathcal{P})$ and let (T, B) be a clique tree of H . Let $e = \{u, v\}$ be any vertex in $\text{LG}(\mathcal{P})$. Then, there does not exist a node $x \in V(T)$ where $B(x)$ contains vertices from both $\text{Inc}(u) \setminus e$ and $\text{Inc}(v) \setminus e$.*

Lemma 7. *Let T be a tree in \mathcal{P} and suppose F is a minimal cut of $G(\mathcal{P})$ that contains precisely one edge e of T . Then, the edges of the two subtrees of $T - e$ are in different connected components of $G(\mathcal{P}) - F$.*

Proof. Let $e = \{u, v\}$. For each $x \in e$, let T_x denote the subtree containing vertex x in $T - e$. For each vertex $x \in e$, all the edges of T_x are in the same connected component of $G(\mathcal{P}) - F$ as x , because e is the only edge of T in F . Since F is a minimal cut of $G(\mathcal{P})$, the endpoints of e are in different connected components of $G(\mathcal{P}) - F$. Hence, the edges of T_u and T_v are also in different connected components of $G(\mathcal{P}) - F$. \square

Characterizing compatibility via cuts

A set \mathcal{F} of cuts of $G(\mathcal{P})$ is *complete* if, for every input tree

$T \in \mathcal{P}$ and every internal edge e of T , there is a cut $F \in \mathcal{F}$ where e is the only edge of T in F .

Lemma 8. $G(\mathcal{P})$ has a complete set of pairwise parallel nice minimal cuts if and only if it has a complete set of pairwise parallel legal minimal cuts.

Proof. The “only if part” follows from the definition of a nice cut. Let \mathcal{F} be a complete set of pairwise parallel legal minimal cuts. Consider any minimal subset \mathcal{F}' of \mathcal{F} that is also complete. Let F be a legal minimal cut of \mathcal{F}' . Since \mathcal{F}' is minimal, there exists an edge $e \in F$ of some input tree T such that e is the only edge of T in F . Also, since e is an internal edge, both subtrees of $T - e$ have at least one edge each. Thus by Lemma 7, both connected components of $G(\mathcal{P}) - F$ have at least one edge each. Hence, F is a nice minimal cut of $G(\mathcal{P})$. It follows that \mathcal{F}' is a complete set of pairwise parallel nice minimal cuts of $G(\mathcal{P})$. \square

We now characterize the compatibility of a profile in terms of minimal cuts in the display graph of the profile.

Theorem 6. A profile \mathcal{P} of unrooted phylogenetic trees is compatible if and only if there exists a complete set of pairwise parallel legal minimal cuts for $G(\mathcal{P})$.

Example 1. For the display graph of Figure 1, let $\mathcal{F} = \{F_1, F_2, F_3, F_4\}$, where $F_1 = \{\{1, 2\}, \{5, 6\}\}$, $F_2 = \{\{2, 3\}, \{6, 7\}, \{5, 6\}\}$, $F_3 = \{\{4, 5\}, \{1, 2\}, \{1, c\}\}$ and $F_4 = \{\{6, 7\}, \{2, f\}\}$. Then, \mathcal{F} is a complete set of pairwise parallel nice minimal cuts.

Theorem 6 has an analog in terms of $LG(\mathcal{P})$. Let us say that a set \mathcal{F} of legal minimal separators of $LG(\mathcal{P})$ is *complete* if for every internal edge e of an input tree T , there exists a separator $F \in \mathcal{F}$ where e is the only vertex of $LG(T)$ in F .

Theorem 7. A profile \mathcal{P} of unrooted phylogenetic trees is compatible if and only if there exists a complete set of pairwise parallel legal minimal separators for $LG(\mathcal{P})$.

This result is a direct consequence of Theorem 6 and Lemmas 2, 5, and 8, so we omit its proof. Instead, we focus on the proof of Theorem 6, for which we need the next fact.

Lemma 9. The following two statements are equivalent.

1. There exists a maximal set \mathcal{F} of pairwise parallel minimal separators of $LG(\mathcal{P})$ where every separator in \mathcal{F} is legal.
2. There exists a complete set of pairwise parallel nice minimal cuts for $G(\mathcal{P})$.

Proof. (i) \Rightarrow (ii): We show that for every internal edge $e = \{u, v\}$ of an input tree T there exists a minimal separator in \mathcal{F} that contains only vertex e from $LG(T)$. Then it follows from Lemmas 2 and 5 that \mathcal{F} is a complete set of pairwise parallel nice minimal cuts for $G(\mathcal{P})$.

As shown in the proof of Theorem 5, $LG(\mathcal{P})_{\mathcal{F}}$ is a restricted minimal triangulation of $LG(\mathcal{P})$. Let (S, B) be a clique tree of $LG(\mathcal{P})_{\mathcal{F}}$. By definition, the vertices in each of the sets $\text{Inc}(u)$ and $\text{Inc}(v)$ form a clique in $LG(\mathcal{P})$. Consider any vertex p of S where $\text{Inc}(u) \subseteq B(p)$ and any vertex q of S where $\text{Inc}(v) \subseteq B(q)$. (Since (S, B) is a clique tree of $LG(\mathcal{P})_{\mathcal{F}}$, such vertices p and q must exist.) Also, by Lemma 6, $p \neq q$, $B(p) \cap (\text{Inc}(v) \setminus \{e\}) = \emptyset$ and $B(q) \cap (\text{Inc}(u) \setminus \{e\}) = \emptyset$.

Let $\pi = p, x_1, x_2, \dots, x_m, q$ be the path from p to q in S where $m \geq 0$. Let $x_0 = p$ and $x_{m+1} = q$. Let x_i be the vertex nearest to p in path π where $i \in [m+1]$ and $B(x_i) \cap (\text{Inc}(u) \setminus \{e\}) = \emptyset$. Let $F = B(x_{i-1}) \cap B(x_i)$. Then by Theorem 2, $F \in \mathcal{F}$. Since $\text{Inc}(u) \cap \text{Inc}(v) = \{e\}$, by the coherence property, $e \in B(x_j)$ for every $j \in [m]$. Thus, $e \in F$. By Lemma 6, $B(x_{i-1}) \cap (\text{Inc}(v) \setminus \{e\}) = \emptyset$. Since $B(x_i) \cap (\text{Inc}(u) \setminus \{e\}) = \emptyset$, $F \cap \text{Inc}(u) = \{e\}$ and $F \cap \text{Inc}(v) = \{e\}$. Thus, for every vertex $e' \in LG(T)$ where $e \neq e'$ and $e \cap e' \neq \emptyset$, $e' \notin F$. Also, since every separator in \mathcal{F} is legal, we have $f \notin F$ for every vertex $f \in LG(T)$ where $f \cap e = \emptyset$. Thus, e is the only vertex of $LG(T)$ in F .

(i) \Leftarrow (ii): Consider any complete set of pairwise parallel nice minimal cuts \mathcal{F}' of $G(\mathcal{P})$. By Lemmas 2 and 5, \mathcal{F}' is a set of pairwise parallel legal minimal separators of $LG(\mathcal{P})$. There exists a maximal set \mathcal{F} of pairwise parallel minimal separators where $\mathcal{F}' \subseteq \mathcal{F}$.

Assume that $\mathcal{F} \setminus \mathcal{F}'$ contains a minimal separator F that is not legal. Then, there must exist a tree $T \in \mathcal{P}$ where at least two nonincident edges $e_1 = \{x, y\}$ and $e_2 = \{x', y'\}$ of T are in F . Consider any internal edge e_3 in T where e_1 and e_2 are in different components of $T - e_3$. Such an edge exists because e_1 and e_2 are nonincident. Since \mathcal{F}' is complete, there exists a cut $F' \in \mathcal{F}'$ where e_3 is the only edge of T in F' . Since F and F' are in \mathcal{F} , they are parallel to each other and vertices e_1 and e_2 are in the same connected component of $LG(\mathcal{P}) - F'$. Thus, by Observation 1, there exists a path between vertices x and x' in $G(\mathcal{P}) - F'$ and edges e_1 and e_2 are also in the same connected component of $G(\mathcal{P}) - F'$. But by Lemma 7 that is impossible.

Thus, every separator of $\mathcal{F} \setminus \mathcal{F}'$ is legal and \mathcal{F} is a maximal set of pairwise minimal separators of $LG(\mathcal{P})$ where every separator in \mathcal{F} is legal. \square

Proof of Theorem 6. By Theorem 5 and Lemma 9, profile \mathcal{P} is compatible if and only if there exists a complete set of pairwise parallel nice minimal cuts for $G(\mathcal{P})$. The rest follows from Lemma 8. \square

Splits and cuts

We first argue that for every nice minimal cut of $G(\mathcal{P})$ we can derive a split of $\mathcal{L}(\mathcal{P})$. We use the following notation: if H is a subgraph of $G(\mathcal{P})$, then $\mathcal{L}(H)$ represents the set of all leaves of H .

Lemma 10. *Let F be a nice minimal cut of $G(\mathcal{P})$ and let G_1 and G_2 be the two connected components of $G(\mathcal{P}) - F$. Then, $\mathcal{L}(G_i) \neq \emptyset$ for $i \in \{1, 2\}$. In particular, $\mathcal{L}(G_1)|\mathcal{L}(G_2)$ is a split of $\mathcal{L}(\mathcal{P})$.*

Proof. Consider G_i for each $i \in \{1, 2\}$. We show that $\mathcal{L}(G_i)$ is non-empty. Since F is nice, G_i contains at least one edge e of $G(\mathcal{P})$. If e is a non-internal edge, then $\mathcal{L}(G_i)$ is non-empty. Assume that $e = \{u, v\}$ is an internal edge of some input tree T . If F does not contain an edge of T , then $\mathcal{L}(T) \subseteq \mathcal{L}(G_i)$ and thus $\mathcal{L}(G_i)$ is non-empty. Assume that F contains one or more edges of T . Let T_u, T_v be the two subtrees of $T - e$. Since F is a nice minimal cut, F contains edges from either T_u or T_v but not both. Without loss of generality assume that F does not contain edges from T_u . Then, every edge of T_u is in the same component as e . Since T_u contains at least one leaf, $\mathcal{L}(G_i)$ is non-empty. Thus, $\mathcal{L}(G_1)|\mathcal{L}(G_2)$ is a split of $\mathcal{L}(\mathcal{P})$. \square

Let $\sigma(F)$ denote the split of $\mathcal{L}(\mathcal{P})$ induced by a nice minimal cut F . If \mathcal{F} is a set of nice minimal cuts of $G(\mathcal{P})$, $\Sigma(\mathcal{F})$ denotes the set of all the non-trivial splits in $\bigcup_{F \in \mathcal{F}} \sigma(F)$. The following result expresses the relationship between complete sets of nice minimal cuts and the compatibility of splits.

Theorem 8. *If $G(\mathcal{P})$ has a complete set of pairwise parallel nice minimal cuts \mathcal{F} , then $\Sigma(\mathcal{F})$ is compatible and any compatible tree for $\Sigma(\mathcal{F})$ is also a compatible tree for \mathcal{P} .*

Example 2. *For the complete set of pairwise parallel nice minimal cuts $\mathcal{F} = \{F_1, F_2, F_3, F_4\}$ for the display graph of Example 1, we have $\sigma(F_1) = abc|defg$, $\sigma(F_2) = abcfg|de$, $\sigma(F_3) = ab|cdefg$, and $\sigma(F_4) = abcde|fg$. Note that these splits are pairwise compatible.*

The proof of Theorem 8 uses the following lemma.

Lemma 11. *Let F_1 and F_2 be two parallel nice minimal cuts of $G(\mathcal{P})$. Then, $\sigma(F_1)$ and $\sigma(F_2)$ are compatible.*

Proof. Let $\sigma(F_1) = U_1|U_2$ and $\sigma(F_2) = V_1|V_2$. Assume that $\sigma(F_1)$ and $\sigma(F_2)$ are incompatible. Thus, $U_i \cap V_j \neq \emptyset$ for every $i, j \in \{1, 2\}$. Let $a \in U_1 \cap V_1$, $b \in U_1 \cap V_2$, $c \in U_2 \cap V_1$ and $d \in U_2 \cap V_2$. Since $\{a, b\} \subseteq U_1$, there exists a path π_1 between leaves a and b in $G(\mathcal{P}) - F_1$. But a and b are in different components of $G(\mathcal{P}) - F_2$.

Thus, an edge e_1 of path π_1 is in the cut F_2 . Similarly, $\{c, d\} \subseteq U_2$ and there exists a path π_2 between labels c and d in $G(\mathcal{P}) - F_1$. Since c and d are in different components of $G(\mathcal{P}) - F_2$, cut F_2 contains an edge e_2 of path π_2 . But π_1 and π_2 are in different components of $G(\mathcal{P}) - F_1$, so edges e_1 and e_2 are in different components of $G(\mathcal{P}) - F_1$. Since $\{e_1, e_2\} \subseteq F_2$, the cuts F_1 and F_2 are not parallel, a contradiction. \square

Proof of Theorem 8. The compatibility of $\Sigma(\mathcal{F})$ follows from Lemma 11 and Theorem 1. Let S be a compatible tree for $\Sigma(\mathcal{F})$, let T be an input tree of \mathcal{P} , let $S' = S|_{\mathcal{L}(T)}$, and let e be any internal edge of T . We show that S' displays $\sigma(e)$.

Let $\sigma(e) = A|B$. There exists a cut $F \in \mathcal{F}$ where e is the only edge of T in F . By Lemma 7, since F is minimal, the leaves of sets A and B are in different components of $G(\mathcal{P}) - F$. Thus, if $\sigma(F) = A'|B'$ then, up to renaming of sets, we have $A \subseteq A'$ and $B \subseteq B'$. Because S displays $\sigma(F)$, S' also displays $\sigma(e)$. Since S' displays all the splits of T , T can be obtained from S' by contracting zero or more edges [10]. Thus, S displays T . Since S displays every tree in \mathcal{P} , S is a compatible tree for \mathcal{P} . \square

Characterizing agreement via cuts

The following characterization of agreement is similar to the one for tree compatibility given by Theorem 6, except for an additional restriction on the minimal cuts.

Theorem 9. *A profile \mathcal{P} has an agreement supertree if and only if $G(\mathcal{P})$ has a complete set \mathcal{F} of pairwise parallel legal minimal cuts where, for every cut $F \in \mathcal{F}$ and for every $T \in \mathcal{P}$, there is at most one edge of T in F .*

Example 3. *One can verify that the display graph of Figure 1 does not meet the conditions of Theorem 9 and, thus, the associated profile does not have an AST. On the other hand, for the display graph of Figure 2, let $\mathcal{F} = \{F_1, F_2, F_3\}$, where $F_1 = \{\{1, 2\}, \{4, 5\}\}$, $F_2 = \{\{1, 2\}, \{5, 6\}\}$ and $F_3 = \{\{2, 3\}, \{6, d\}\}$. For any given input tree T , every cut in \mathcal{F} has at most one edge of T . Also, \mathcal{F} is a complete set of pairwise parallel legal minimal cuts. Thus, by Theorem 9, the input trees of Figure 2 have an AST.*

The analogue of Theorem 9 for $LG(\mathcal{P})$ stated next follows from Theorem 9 and Lemmas 2, 5, and 8.

Theorem 10. *A profile \mathcal{P} has an agreement supertree if and only if $LG(\mathcal{P})$ has a complete set \mathcal{F} of pairwise parallel legal minimal separators where, for every $F \in \mathcal{F}$ and every $T \in \mathcal{P}$, there is at most one vertex of $LG(T)$ in F .*

Theorem 9 follows from Lemma 8 and the next result.

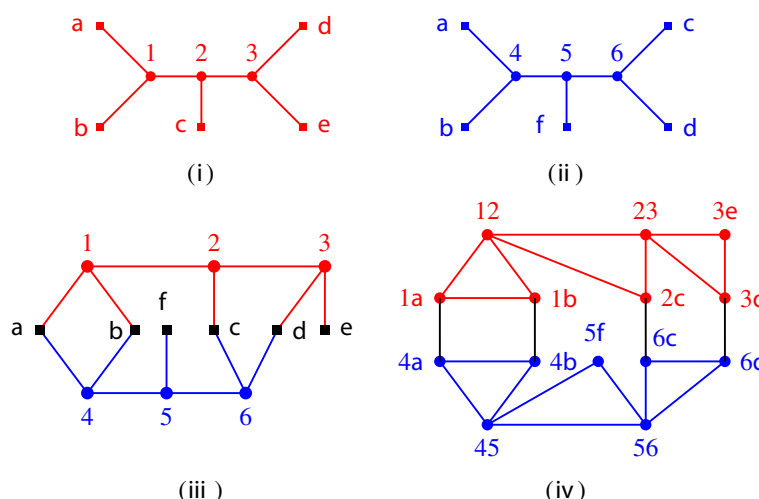


Figure 2 Agreeing trees. (i) First input tree. (ii) Second input tree, which agrees with the first. (iii) Display graph of the input trees. (iv) Edge label intersection graph of the input trees, where label uv represents edge $\{u, v\}$ of the display graph.

Lemma 12. *A profile \mathcal{P} has an agreement supertree if and only if $G(\mathcal{P})$ has a complete set \mathcal{F} of pairwise parallel nice minimal cuts where, for every cut $F \in \mathcal{F}$ and every $T \in \mathcal{P}$, there is at most one edge of T in F .*

The rest of the section is devoted to the proof of Lemma 12.

Let S be an AST of \mathcal{P} and let $e = \{u, v\}$ be an edge of S . Let S_u and S_v be the subtrees of $S - e$ containing u and v , respectively. Let $L_u = \mathcal{L}(S_u)$ and $L_v = \mathcal{L}(S_v)$. Thus, $\sigma_e(S) = L_u | L_v$. Assume that there exists an input tree T where $\mathcal{L}(T) \cap L_x \neq \emptyset$ for each $x \in \{u, v\}$. Then there exists an edge $f \in E(T)$ where, if $\sigma_f(T) = A_1 | A_2$, then $A_1 \subseteq L_u$ and $A_2 \subseteq L_v$. (If there were no such edge, $S|_{\mathcal{L}(T)}$ would contain a split that is not in T and would thus not be isomorphic to T .) We call e an *agreement edge* of S corresponding to edge f of T . Note that there does not exist any other edge f' of T where e is also an agreement edge of S with respect to edge f' of T .

The *cut function* of an AST S of \mathcal{P} is the mapping Ψ from $E(S)$ to subsets of edges of $G(\mathcal{P})$ defined as follows. For every $e \in E(S)$, an edge f of an input tree T is in $\Psi(e)$ if and only if e is an agreement edge of S corresponding to edge f of T . Observe that Ψ is uniquely defined. Given an edge $e \in E(S)$, we define a set V_x for each $x \in e$ as follows. For every $T \in \mathcal{P}$, let $V_{x,T}$ consist of all the vertices of the minimal subtree of T connecting the labels in $\mathcal{L}(T) \cap L_x$. Then, $V_x = \bigcup_{T \in \mathcal{P}} V_{x,T}$. Note that if $e = \{u, v\}$ then $\{V_u, V_v\}$ is a partition of $V(G(\mathcal{P}))$.

Lemma 13. *Let S be an AST of \mathcal{P} and let Ψ be the cut function of S . Then, for every edge $e \in E(S)$,*

- (i) $\Psi(e)$ is a cut of $G(\mathcal{P})$ and
- (ii) $\Psi(e)$ is a minimal cut of $G(\mathcal{P})$ if and only if $G(\mathcal{P}) - \Psi(e)$ has exactly two connected components.

Proof. (i) Let $e = \{u, v\}$. We show that $G(\mathcal{P}) - \Psi(e)$ does not contain an edge whose endpoints are in distinct sets of $\{V_u, V_v\}$. Assume the contrary. Let $f = \{x, y\}$ be an edge of $G(\mathcal{P}) - \Psi(e)$ where $x \in V_u$ and $y \in V_v$.

Since $f \in G(\mathcal{P}) - \Psi(e)$, $f \notin \Psi(e)$. Suppose f is an edge of input tree T . There are two cases.

1. $\Psi(e)$ does not contain an edge of T . Then, there exists an endpoint p of e where $\mathcal{L}(T) \subseteq L_p$. Without loss of generality, let $u = p$. Then, $V(T) \subseteq V_u$ and thus $y \in V_u$, a contradiction.
2. $\Psi(e)$ contains an edge $f' \neq f$ of T . Let $f' = \{r, s\}$ and let $L_r \subseteq L_u$ and $L_s \subseteq L_v$. Let x, r be the vertices of f and f' where $L_x \subset L_r$. Since T is a phylogenetic tree, such vertices x and r exist. Since $L_r \subseteq L_u$, both the endpoints of f are in V_u , a contradiction.

Thus, $G(\mathcal{P}) - \Psi(e)$ does not contain an edge whose endpoints are in different sets of $\{V_u, V_v\}$. Since V_u and V_v are non-empty, $\Psi(e)$ is a cut of $G(\mathcal{P})$.

(ii) The “only if” part follows from the definition of a minimal cut. We now prove the “if” part. Let $e = \{u, v\}$. Assume that $G(\mathcal{P}) - \Psi(e)$ has exactly two connected components. From the proof of (i), V_u and V_v are the vertex sets of those two connected components. Consider any edge $f \in \Psi(e)$. The endpoints of f are in different sets of $\{V_u, V_v\}$ and thus are in different connected components of $G(\mathcal{P}) - \Psi(e)$. Hence, $G(\mathcal{P}) - (\Psi(e) \setminus \{f\})$ is connected. Thus, if $G(\mathcal{P}) - \Psi(e)$ has exactly two connected components, $\Psi(e)$ is a minimal cut of $G(\mathcal{P})$. \square

The next observation summarizes two basic facts about cut functions.

Observation 2. Let S be an AST of \mathcal{P} . Then, the cut function Ψ of S has the following properties.

1. For any two distinct edges e_1 and e_2 in $E(S)$, $\Psi(e_1) \neq \Psi(e_2)$.
2. Let $e = \{u, v\}$ be an edge of S . For any input tree T where $\mathcal{L}(T) \cap L_v \neq \emptyset$, all the labels of $\mathcal{L}(T) \cap L_v$ are in the same connected component of $G(\mathcal{P}) - \Psi(e)$.

Let S be an AST of \mathcal{P} and let e be an edge of S . Although Lemma 13 shows that $\Psi(e)$ is a cut of $G(\mathcal{P})$, $\Psi(e)$ may not be minimal. We now argue that we can always construct an agreement supertree whose cut function gives minimal cuts.

Lemma 14. If \mathcal{P} has an AST, then it has an AST S of \mathcal{P} whose cut function Ψ satisfies the following: For every edge $e \in S$, $\Psi(e)$ is a minimal cut of $G(\mathcal{P})$.

We prove Lemma 14 by arguing that any AST that fails to satisfy the required cut minimality property can be transformed into one that does, through repeated application of the “splitting” operation, defined next.

Suppose $e = (u, v)$ is an edge of S where $\Psi(e)$ is not minimal. Let $\{L_1, \dots, L_m\}$ be the partition of L_v where for every $i \in [m]$, $L_i = \mathcal{L}(C) \cap L_v$ for some connected component C in $G(\mathcal{P}) - \Psi(e)$. We assume without loss of generality that $m > 1$ (if not, we can just exchange the roles of u and v). Let R_v be the rooted tree derived from S_v by distinguishing vertex v as the root. Let $R_{v,i}$ be the (rooted) tree obtained from the minimal subtree of R_v connecting the labels in L_i by distinguishing the vertex closest to v as the root and suppressing every other vertex that has degree two. To split edge e at u is to construct a new tree S' from S in two steps: (i) delete the vertices of R_v from S and (ii) for every $i \in [m]$, add an edge from u to the root of $R_{v,i}$.

Observation 3. Let S be an AST of \mathcal{P} and let Ψ be the cut function of S . Let S' be the tree derived by splitting edge $e = \{u, v\}$ at u . Consider any connected component C of $G(\mathcal{P}) - \Psi(e)$ where $\mathcal{L}(C) \cap L_v \neq \emptyset$. Then, for every $X \subseteq (\mathcal{L}(C) \cap L_v)$, $S|_X$ and $S'|_X$ are isomorphic.

The next observation follows from the definition of AST.

Observation 4. Let S and T be two phylogenetic trees where $\mathcal{L}(T) \subseteq \mathcal{L}(S)$ and T agrees with S . Then, T and $S|_U$ agree for every U such that $\mathcal{L}(T) \subseteq U \subseteq \mathcal{L}(S)$.

Lemma 15. Let S be an AST of \mathcal{P} and let $e = \{u, v\}$ be an edge of S . Let S' be the tree derived by splitting edge e at u . Then, S' is an AST of \mathcal{P} .

Proof. By construction, S' is a phylogenetic tree over $\mathcal{L}(\mathcal{P})$. As before, let $\{L_1, \dots, L_m\}$ be the partition of L_v where for every $i \in [m]$, $L_i = \mathcal{L}(C) \cap L_v$ for some connected component C in $G(\mathcal{P}) - \Psi(e)$. Consider any input tree T of profile \mathcal{P} . We prove that T and S' agree. There are three cases.

Case 1: $\mathcal{L}(T) \subseteq L_u$. Since $\mathcal{L}(T) \subseteq L_u$, by Observation 4, T and $S|_{L_u}$ agree. By the definition of the split operation, trees $S|_{L_u}$ and $S'|_{L_u}$ are isomorphic. Thus, T and S' agree.

Case 2: $\mathcal{L}(T) \subseteq L_v$. By Observation 2(ii), $\mathcal{L}(T) \subseteq L_i$ for some $i \in [m]$. Since T and S agree and $\mathcal{L}(T) \subseteq L_i$, by Observation 4, T and $S|_{L_i}$ agree. By construction, trees $S|_{L_i}$ and $S'|_{L_i}$ are isomorphic. Thus, T and S' agree.

Case 3: $(\mathcal{L}(T) \cap L_u \neq \emptyset)$ and $(\mathcal{L}(T) \cap L_v \neq \emptyset)$. By Observation 2(ii), $\mathcal{L}(T) \cap L_v \subseteq L_i$ for some $i \in [m]$. Since T and S agree and $\mathcal{L}(T) \subseteq (L_u \cup L_i)$, by Observation 4, T also agrees with $S|_{(L_u \cup L_i)}$. By construction, trees $S|_{(L_u \cup L_i)}$ and $S'|_{(L_u \cup L_i)}$ are isomorphic. Thus, T and S' agree. It follows that T and S' agree.

Thus, S' is an AST of \mathcal{P} . \square

Observe that if S' is the tree obtained by splitting edge $e = \{u, v\}$ of S at u , then the edges of $E(S_u)$ are in both S and S' .

Therefore, $E(S) \setminus E(S_u) = E(S') \setminus E(S_u)$ and $E(S') \setminus E(S_u) = E(S') \setminus E(S)$.

Lemma 16. Let S be an AST of \mathcal{P} and let $e = \{u, v\}$ be an edge of S . Let S' be the tree obtained by splitting e at u . Let Ψ, Ψ' be the cut functions of S and S' respectively. Consider any edge $f \in E(S') \setminus E(S)$. There exists an edge $e' \in E(S) \setminus E(S')$ where $\Psi'(f) \subseteq \Psi(e')$. Furthermore, if $\Psi(e')$ is a minimal cut of $G(\mathcal{P})$ then $\Psi'(f) = \Psi(e')$ and $\Psi'(f)$ is a minimal cut of $G(\mathcal{P})$.

Proof. Let $f = \{x, y\}$ and let x be the vertex of f where $L_x \subseteq L_v$. Let S_p be the minimal subtree of S connecting the labels in L_x . Let p be the vertex of S_p closest to u in S . Let q be the vertex adjacent to p in the path from p to u . Let $e' = \{p, q\}$. Note that, $L_x \subseteq L_p$. Since $L_x \subseteq L_v$, e' is an edge of $E(S) \setminus E(S')$. Consider any tree T that has an edge f_1 in $\Psi'(f)$. We show that $\mathcal{L}(T) \cap L_x = \mathcal{L}(T) \cap L_p$. It then follows that $f_1 \in \Psi(e')$ and thus, $\Psi'(f) \subseteq \Psi(e')$.

Since $L_x \subseteq L_p$, $(L_x \cap \mathcal{L}(T)) \subseteq (\mathcal{L}(T) \cap L_p)$. By Observation 2(ii), all the labels in $\mathcal{L}(T) \cap L_v$ are in the same connected component of $G(\mathcal{P}) - \Psi(e)$. Thus, all the labels in $L_x \cup (L_p \cap \mathcal{L}(T))$ are in the same connected of $G(\mathcal{P}) - \Psi(e)$. If $(L_p \cap \mathcal{L}(T)) \not\subseteq (L_x \cap \mathcal{L}(T))$, then $S|_{(L_x \cup (L_p \cap \mathcal{L}(T)))}$ and

$S'_{|(L_x \cup (L_p \cap \mathcal{L}(T)))}$ are not isomorphic, contradicting Observation 3. Thus, $(L_p \cap \mathcal{L}(T)) \subseteq (L_x \cap \mathcal{L}(T))$.

Assume that $\Psi(e')$ is a minimal cut of $G(\mathcal{P})$. Then, all the labels in L_p are in the same connected component of $G(\mathcal{P}) - \Psi(e')$. By Observation 3, $L_p = L_x$. Thus, $\Psi'(f)$ is also a minimal cut of $G(\mathcal{P})$. \square

Lemma 17. *Let S be an AST of \mathcal{P} and Ψ be the cut function of S . Let E_0 be the set of all edges e of S such that $\Psi(e)$ is not a minimal cut of $G(\mathcal{P})$. Choose any edge $e^* = \{u, v\} \in E_0$ such that $|\Psi(e^*)| = \max_{e \in E_0} |\Psi(e)|$. Let S' be the tree obtained from S by splitting e^* at u and let Ψ' be the cut function of S' . We have the following.*

1. For any edge $f \in E(S')$, if $|\Psi'(f)| > |\Psi(e^*)|$ then $\Psi'(f)$ is a minimal cut of $G(\mathcal{P})$.
2. Let P be the set of all edges x in S such that $|\Psi(e^*)| = |\Psi(x)|$ and $\Psi(x)$ is not a minimal cut. Let P' be the set of all edges x in S' such that $|\Psi(e^*)| = |\Psi(x)|$ and $\Psi'(x)$ is not a minimal cut. Then, $|P'| < |P|$.

Proof. (i) Consider any edge $f \in E(S')$ where $|\Psi'(f)| > |\Psi(e^*)|$. If $f \in E(S) \cap E(S')$, then $\Psi(f) = \Psi'(f)$. Since $|\Psi(f)| > |\Psi(e^*)|$, by assumption $\Psi(f)$ is a minimal cut of $G(\mathcal{P})$. Thus, $\Psi'(f)$ is also a minimal cut of $G(\mathcal{P})$. Assume that $f \in E(S') \setminus E(S)$. By Lemma 16, there exists an edge $e' \in E(S)$ where $\Psi'(f) \subseteq \Psi(e')$. Since $|\Psi'(f)| > |\Psi(e^*)|$, $|\Psi(e')| > |\Psi(e^*)|$. Thus, by assumption $\Psi(e')$ is a minimal cut of $G(\mathcal{P})$. From Lemma 16, it follows that $\Psi(e') = \Psi'(f)$ and $\Psi'(f)$ is a minimal cut of $G(\mathcal{P})$.

(ii) Let $Q = P \cap (E(S) \setminus E(S'))$ and $Q' = P' \cap (E(S') \setminus E(S))$. It suffices to show that $|Q'| < |Q|$. Consider any edge $f \in Q'$. By Lemma 16, there exists an edge $e' \in E(S) \setminus E(S')$ where $\Psi'(f) \subseteq \Psi(e')$. Thus, $|\Psi(e')| \geq |\Psi'(f)|$. If $|\Psi(e')| > |\Psi'(f)|$, then by assumption $\Psi(e')$ is a minimal cut and thus by Lemma 16 $|\Psi(e')| = |\Psi'(f)|$, a contradiction.

Thus, $\Psi(e') = \Psi'(f)$. Also, since $\Psi'(f)$ is not a minimal cut, by Lemma 16, neither is $\Psi(e')$. If $e' = e^*$, then all vertices of V_v are in the same connected component of $G(\mathcal{P}) - \Psi(e)$, contradicting the assumption that it is possible to split e^* at u . Thus, $e' \neq e^*$. Hence, we can conclude that for every edge $f \in Q'$, there exists an edge $e' \in (Q \setminus \{e^*\})$, where $\Psi'(f) = \Psi(e')$.

Let f_1 and f_2 be any two distinct edges in Q' . Let e_1 and e_2 be the edges of $Q \setminus \{e^*\}$ where $\Psi'(f_1) = \Psi(e_1)$ and $\Psi'(f_2) = \Psi(e_2)$. If $e_1 = e_2$, then $\Psi'(f_1) = \Psi'(f_2)$, contradicting Observation 2(i). Thus, $e_1 \neq e_2$. Since $e \in Q$ and $e \notin Q'$, it follows that $|Q'| \leq |Q| - 1$, and thus $|Q'| < |Q|$. \square

Proof of Lemma 14. Let S be an AST of \mathcal{P} and Ψ be the cut function of S . Do the following while S contains an edge e such that $\Psi(f)$ is not a minimal cut of $G(\mathcal{P})$:

Pick an edge e^* satisfying the conditions of Lemma 17, and apply a split operation at e^* ; let S' be the resulting tree. By Lemma 15, S' is also an AST of \mathcal{P} . Let Ψ' be the cut function of S' . Set S to S' and Ψ to Ψ' .

We only need to prove that the total number of iterations, s , is finite. An AST of \mathcal{P} has at most $2|\mathcal{L}(\mathcal{P})|$ vertices. Also, $|\Psi(e)| \geq 1$ for any edge e of S . It thus follows from Lemma 17 that s is finite. \square

Proof of Lemma 12. (\Leftarrow) Assume that \mathcal{P} has an AST. Then, by Lemma 14, \mathcal{P} has an AST S whose cut function Ψ has the property that, for every edge $e \in E(S)$, $\Psi(e)$ is a minimal cut of $G(\mathcal{P})$. Let \mathcal{F} be the set of all $\Psi(e)$ such that e is an internal edge of S . Then, \mathcal{F} is a set of minimal cuts of $G(\mathcal{P})$. Further, by definition of Ψ , for every $F \in \mathcal{F}$ and for every $T \in \mathcal{P}$, F contains at most one edge of T . Thus every cut in \mathcal{F} is legal. We now prove that \mathcal{F} is a complete set of pairwise parallel nice minimal cuts of $G(\mathcal{P})$.

We first argue that every cut in \mathcal{F} is nice. Consider any $F \in \mathcal{F}$. Let $e = \{u, v\}$ be the internal edge of S where $\Psi(e) = F$. Let T be an input tree that has an internal edge f in $\Psi(e)$. Since e is an internal edge at least one such input tree exists; otherwise $\Psi(e)$ is not a minimal cut. Now, by definition, f is the only edge of T in $\Psi(e)$, so, by Lemma 7, each of the two connected components of $G(\mathcal{P}) - \Psi(e)$ has at least one non-internal edge of T . Hence, F is a nice minimal cut of $G(\mathcal{P})$.

To prove that the cuts in \mathcal{F} are pairwise parallel, we argue that for any two distinct internal edges e_1 and e_2 of S , $\Psi(e_1)$ and $\Psi(e_2)$ are parallel. There exist vertices $x \in e_1$ and $y \in e_2$ where $L_x \subseteq L_y$. For every edge $f \in \Psi(e_1)$, we show that $f \in \Psi(e_2)$ or $f \subseteq V_y$. It then follows that $\Psi(e_1)$ and $\Psi(e_2)$ are parallel. Let f be an edge of input tree T . Then there exists $z \in f$ where $L_z \subseteq L_x$. Thus, $L_z \subseteq L_y$ and $z \in V_y$. By Lemma 13, all the vertices of V_y are in the same connected component of $G(\mathcal{P}) - \Psi(e_2)$. Thus, $f \in \Psi(e_2)$ or $f \subseteq V_y$.

Lastly, we show that \mathcal{F} is complete. Consider any internal edge $f = \{p, q\}$ of some input tree T . Since S is an AST of \mathcal{P} , there exists an edge $e = \{u, v\}$ where, up to relabeling of sets, $L_p \subseteq L_u$ and $L_q \subseteq L_v$. Thus, e is an agreement edge of S corresponding to f , so $f \in \Psi(e)$. Since f is an internal edge, e is also an internal edge of S and thus $\Psi(e) \in \mathcal{F}$. Hence, for every internal edge f of an input tree there is a cut $F \in \mathcal{F}$ where $f \in F$. Thus, \mathcal{F} is complete.

(\Rightarrow) Assume that there exists a complete set \mathcal{F} of pairwise parallel nice minimal cuts of $G(\mathcal{P})$ where, for every $F \in \mathcal{F}$ and every $T \in \mathcal{P}$, F contains at most one edge of T . By Theorem 8, $\Sigma(\mathcal{F})$ is compatible and, by Theorem 1, there exists an unrooted tree S where $\Sigma(\mathcal{F}) = \Sigma(S)$. We prove that S is an AST of \mathcal{P} by showing that $\Sigma(S_{|\mathcal{L}(T)}) = \Sigma(T)$ for every input tree $T \in \mathcal{P}$.

Consider an input tree T of \mathcal{P} . Let $X_1|X_2$ be the non-trivial split of T corresponding to edge $f \in E(T)$. Since

\mathcal{F} is complete, there exists a cut $F \in \mathcal{F}$ where $f \in \mathcal{F}$. If $\sigma(F) = Y_1|Y_2$, by Lemma 7, up to relabeling of sets, $X_i \subseteq Y_i$ for every $i \in \{1, 2\}$. Since $\sigma(F)$ is a split of S , this implies that $\Sigma(T) \subseteq \Sigma(S_{|\mathcal{L}(T)})$.

Consider any non-trivial split $P_1|P_2$ of $\Sigma(S)$ where $P_i \cap \mathcal{L}(T) \neq \emptyset$ for each $i \in \{1, 2\}$. Let $Q_i = P_i \cap \mathcal{L}(T)$ for each $i \in \{1, 2\}$. Since $\Sigma(S) = \Sigma(\mathcal{F})$, there exists a cut $F \in \mathcal{F}$ where $\sigma(F) = P_1|P_2$. Since P_1 and P_2 are in different connected components of $G(\mathcal{P}) - F$, Q_1 and Q_2 are also in different connected components of $G(\mathcal{P}) - F$. Thus, F contains an edge f' of T . Since F does not contain any other edge of T , $\sigma(f') = Q_1|Q_2$. Thus, $\Sigma(S_{|\mathcal{L}(T)}) \subseteq \Sigma(T)$. \square

Relationship to legal triangulations

Taken together, Theorems 3 and 6 say that $G(\mathcal{P})$ has a complete set of pairwise parallel legal minimal cuts if and only if it has a legal triangulation. The connection between legal triangulations and complete sets of pairwise parallel legal minimal cuts is through the existence (or nonexistence) of a compatible tree. Here we make the connection explicit, showing how, from a set of pairwise parallel legal minimal cuts, one can construct a legal triangulation of $G(\mathcal{P})$ without going through a compatible tree. We leave the other direction—going from a triangulation to a set of cuts—to the reader.

Let \mathcal{F} be a complete set of pairwise parallel legal minimal cuts of $G(\mathcal{P})$. We assume that the elements of \mathcal{F} are ordered in some arbitrary, but fixed, manner, and that no proper subset of \mathcal{F} is also complete. For each $F \in \mathcal{F}$, we build a pair (X_F, Y_F) where X_F and Y_F are vertex separators of $G(\mathcal{P})$, and $X_F, Y_F \subseteq \{u : u \text{ is the endpoint of some edge in } F\}$. The collection of pairs $\{(X_F, Y_F) : F \in \mathcal{F}\}$ is not unique, as it depends on the order in which \mathcal{F} is arranged. We say that a cut $F \in \mathcal{F}$ *differentiates* an internal edge $e = \{x, y\}$ if $x \in X_F$ and $y \in Y_F$.

For each $F \in \mathcal{F}$, let $F_i = E(T_i) \cap F$ for each $i \in [k]$, and let $\hat{F} \subseteq F$ denote the set of all edges e such that $e \in F_i$ for some $i \in [k]$ with $|F_i| = 1$. Note that if $|F_i| > 1$, all edges in F_i must share a common endpoint. Let A_F and B_F denote the two connected components of $G(\mathcal{P}) - F$.

For each cut F in \mathcal{F} , we build (X_F, Y_F) as follows.

1. For each internal edge $e \in \hat{F}$:
 - (a) If no cut preceding F differentiates e , add $e \cap V(A_F)$ to X_F and $e \cap V(B_F)$ to Y_F .
 - (b) Otherwise, suppose cut $I \in \mathcal{F}$, which precedes F , differentiates e . Let Q be the connected component of $G(\mathcal{P}) - I$ where $E(Q) \cap F \neq \emptyset$. (Note that Q is unique, since I and F are parallel.) Let v be the unique endpoint of e in Q . Add v to X_F and Y_F .

2. For each non-internal edge $e \in \hat{F}$, add the non-leaf endpoint of e to both X_F and Y_F .
3. For each $i \in [k]$ such that $|F_i| > 1$, add the common endpoint of the edges of F_i to both X_F and Y_F .

By construction and the properties of \mathcal{F} , every edge internal edge of $G(\mathcal{P})$ is differentiated by some cut $F \in \mathcal{F}$. Further, the sets X_F and Y_F have the form $X_F = \{x_1, \dots, x_m, z_1, \dots, z_p\}$ and $Y_F = \{y_1, \dots, y_m, z_1, \dots, z_p\}$, where $m > 0$, $p \geq 0$, and for every $i \in [m]$, $\{x_i, y_i\}$ is an internal edge of $G(\mathcal{P})$ that is differentiated by F . Let

$$O_F = \{\{x_1, \dots, x_j, y_j, \dots, y_m, z_1, \dots, z_p\} : j \in [m]\}.$$

We now state how to go from a complete set of pairwise parallel legal cuts to a legal triangulation. As in Section ‘Preliminaries’, given a graph G and a collection Δ of subsets of $V(G)$, G_Δ denotes the graph derived from G by making the set of vertices of X a clique for every $X \in \Delta$.

Theorem 11. *Let Δ be the collection of subsets of $V(G(\mathcal{P}))$ given by*

$$\Delta = \{N_{G(\mathcal{P})}(\ell) : \ell \text{ is a leaf in } G(\mathcal{P})\} \cup \bigcup_{F \in \mathcal{F}} (\{X_F, Y_F\} \cup O_F). \quad (1)$$

Then, $G(\mathcal{P})_\Delta$ is a legal triangulation of $G(\mathcal{P})$.

The proof of Theorem 11 relies on a series of auxiliary lemmas, for which we introduce some new notation. For each $F \in \mathcal{F}$, F_\cup denotes $X_F \cup Y_F$ and F_\cap denotes $X_F \cap Y_F$. Also, we abbreviate $G(\mathcal{P})_\Delta$ to G_Δ , where Δ is the set defined in Equation (1).

Lemma 18. *Let F and I be two distinct cuts of \mathcal{F} , and let x be a vertex of F_\cup . Suppose x lies in the connected component of $G(\mathcal{P}) - I$ that does not contain edges of F . Then, $x \in I_\cap$.*

Proof. Let $E_{F,x}$ be the set of all edges of F that contain x and let $E_{I,x}$ be the set of all edges of I that contain x . We must have $E_{F,x} \subseteq E_{I,x} \subseteq I$. If $|E_{I,x}| > 1$, then $x \in I_\cap$. Thus, assume that $|E_{I,x}| = 1$. Let $E_{I,x} = \{e\}$, where $e = \{x, y\}$. Since $E_{F,x} \subseteq E_{I,x}$ and $|E_{F,x}| \geq 1$, $E_{F,x} = \{e\}$. We can assume that y is not a leaf (since, otherwise, $x \in I_\cap$). Let $E_{I,y}$ be the set of edges of I with y as an endpoint. Vertex y lies in the component of $G - F$ that does not contain I . Thus, every edge in $E_{I,y}$ is also present in F . If $|E_{I,y}| > 1$, then there is more than one edge in F with y as an endpoint and by construction, $x \notin F_\cup$. Hence, $|E_{I,y}| = 1$, and so $E_{I,y} = \{e\}$.

Let J be the cut that differentiates e . If $F = J$ then by construction, $x \in I_\cap$. Thus, assume that $F \neq J$. If J is in the same connected component of $G(\mathcal{P}) - F$ as I , then, by construction $x \notin F_\cup$, which is a contradiction. Thus, J is

in the connected component of $G(\mathcal{P}) - F$ that does not contain I and, by construction, $x \in I_\cap$. \square

Lemma 19. *Let $F \in \mathcal{F}$. For every edge $\{u, v\}$ in G_Δ , (i) if $u \in V(A_F) \setminus F_\cap$, then $v \notin V(B_F) \setminus Y_F$, and (ii) if $u \in V(B_F) \setminus F_\cap$, then $v \notin V(A_F) \setminus X_F$.*

Proof. Without loss of generality, we consider only the case where $u \in V(A_F) \setminus F_\cap$. Suppose that $v \in V(B_F) \setminus Y_F$. If $e \in E(G(\mathcal{P}))$, then $e \in F$ and hence, by construction, at least one of u and v is in F_\cup . But $v \notin Y_F$, so $u \in F_\cap$, a contradiction.

Thus, e must be a fill-in edge. Since $e \not\subseteq F_\cup$, there must be a cut $I \in \mathcal{F}$, $I \neq F$, such that $e \subseteq I_\cup$. If $E(A_F) \cap I \neq \emptyset$, then by Lemma 18, $v \in F_\cap$, a contradiction. Thus, assume that $E(B_F) \cap I \neq \emptyset$. Then, by Lemma 18, $u \in F_\cap$, another contradiction. \square

A clique of G_Δ is *illegal* if it contains a fill-in edge with a leaf as an endpoint or it contains an internal edge along with any another edge of $G(\mathcal{P})$. An illegal clique violates one of the legal triangulation conditions (LT1) or (LT2) stated in Section ‘Display graphs and edge label intersection graphs.’

Lemma 20. *Let F be a cut of \mathcal{F} and let H be the subgraph of G_Δ induced by vertices of F_\cup . Then, H is triangulated and contains no illegal clique.*

Proof. Let $X_F = \{x_1, \dots, x_m, z_1, \dots, z_p\}$ and $Y_F = \{y_1, \dots, y_m, z_1, \dots, z_p\}$, where for every $i \in [m]$, $x_i \in V(A_F)$, $y_i \in V(B_F)$ and $\{x_i, y_i\}$ is an internal edge of $G(\mathcal{P})$. Note that $F_\cap = \{z_1, \dots, z_p\}$.

Claim. *For every $i, j \in [m]$ where $i > j$, $e = \{x_i, y_j\} \notin E(H)$.*

Proof. Assume that $e \in E(H)$. By construction of (X_F, Y_F) , e is a fill-in edge. Since no set in O_F contains both x_i and y_j , there is a cut $I \in \mathcal{F}$ where $e \subseteq I_\cup$. Since F and I are parallel, only one of the two sets $I \cap E(A_F)$ or $I \cap E(B_F)$ is non-empty. Assume that $I \cap E(A_F) \neq \emptyset$. Then by Lemma 18, $y_j \in F_\cap$, a contradiction. Similarly, if $I \cap E(B_F) \neq \emptyset$, then by Lemma 18, $x_i \in F_\cap$, a contradiction. \square

Let C be a chordless cycle of length at least four in H . Since X_F and Y_F are cliques in G_Δ , if C contains more than two vertices from one of X_F or Y_F , then C must contain a chord. Hence, C has exactly four vertices, with exactly two vertices each from X_F and Y_F . We will first show that $z_i \notin C$ for any $i \in [p]$. Assume that $z_i \in C$ for some $i \in [p]$. Then, of the remaining three vertices of C , at least two of them belong to one of X_F and Y_F . Let a, b be those two vertices.

Without loss of generality assume that $\{a, b\} \subseteq X_F$. Since, $F_\cap \subseteq X_F$, vertices z_i, a, b form a clique in H . Thus, C is not chordless, a contradiction.

Let x_i, x_j be the vertices of X_F in C where $1 \leq i < j \leq m$. Similarly, let $y_{i'}, y_{j'}$ be the vertices of Y_F in C where $1 \leq i' < j' \leq m$. Now, either $i \leq i'$ or $i > i'$. If $i \leq i'$, then $\{x_1, \dots, x_i, y_{i'}, \dots, y_m, z_1, \dots, z_p\} \in O_F$ and thus vertices $x_i, y_{i'}, y_{j'}$ form a clique. Hence, C is not chordless, a contradiction. If $i > i'$, then from the above claim neither of the edges $\{x_i, y_{i'}\}$ and $\{x_j, y_{j'}\}$ can exist. Thus, vertex $y_{i'}$ cannot be in C , a contradiction. Hence, H does not contain a chordless cycle and is triangulated.

Assume that H contains an illegal clique H' ; that is, H' contains two internal edges e and e' . By construction, F_\cup cannot contain a leaf. By the legality of F and the construction of F_\cup , edges e and e' are from different input trees and both are differentiated by F . Let $e = \{x_i, y_i\}$ for some $i \in [m]$ and let $e' = \{x_j, y_j\}$ for some $j \in [m]$. Without loss of generality, assume that $i < j$. By the above claim, there is no edge between x_j and y_i in H ; thus, H' is not a clique, a contradiction. \square

Lemma 21. *G_Δ is chordal.*

Proof. Assume the contrary. Let C be a chordless cycle of length at least four in G_Δ . By construction, C cannot contain a leaf. There are two cases.

Case 1: There are vertices $u, v \in V(C)$ and a cut $F \in \mathcal{F}$ where $u \in X_F \setminus F_\cap$ and $v \in Y_F \setminus F_\cap$.

We have two subcases.

- Suppose C contains a vertex $x \in F_\cap$. Then, there exists a path u, x, v in C . Because C is a cycle, there must exist an edge between a vertex $u' \in V(A_F) \setminus x$ and $v' \in V(B_F) \setminus x$. Since C is chordless, $u' \notin F_\cap$ and $v' \notin F_\cap$. Thus, $u' \in V(A_F) \setminus F_\cap$ and $v' \in V(B_F) \setminus F_\cap$. By Lemma 19, if $u' \in V(A_F) \setminus X_F$ then there is no edge between u' and v' . Thus, $u' \in X_F \setminus F_\cap$. Similarly, $v' \in Y_F \setminus F_\cap$. If $u \neq u'$ or $v \neq v'$, C cannot be chordless. Thus, $u = u'$ and $v = v'$ and C has length three, a contradiction.
- Suppose C does not contain a vertex of F_\cap . Since $u \in V(A_F) \setminus F_\cap$, $v \in V(B_F) \setminus F_\cap$ and F is a cut, there must exist two edges $e_1 = \{x_1, y_1\}$ and $e_2 = \{x_2, y_2\}$ in C where $\{x_1, x_2\} \subseteq V(A_F) \setminus F_\cap$ and $\{y_1, y_2\} \subseteq V(B_F) \setminus F_\cap$. If $x_1 \in V(A_F) \setminus X_F$, then by Lemma 19 there cannot exist an edge between x_1 and y_1 . Thus, $x_1 \in X_F \setminus F_\cap$. Similarly, $x_2 \in X_F \setminus F_\cap$ and $\{y_1, y_2\} \subseteq Y_F \setminus F_\cap$. Since X_F and Y_F are cliques in G_Δ , there exist edges $\{x_1, x_2\}$ and $\{y_1, y_2\}$. Thus, there cannot exist any other vertex in C and hence $V(C) \subseteq F_\cup$. But, by Lemma 20 subgraph of G_Δ induced by vertices of F_\cup is triangulated. Thus, C is not chordless, a contradiction.

Case 2: There is no cut $F \in \mathcal{F}$ with vertices $u \in X_F \setminus F_\cap$ and $v \in Y_F \setminus F_\cap$ such that $u, v \in V(C)$. Thus, for every cut $F \in \mathcal{F}$ at most two vertices of $V(C)$ are in F_\cup . Let x_1, x_2, x_3, x_4 be a path of length four in C . For every $i \in \{1, 2, 3\}$, let $F^{(i)} \in \mathcal{F}$ be the cut where $\{x_i, x_{i+1}\} \subseteq F_\cup^{(i)}$. We will first show that such cuts exist and are distinct.

Recall that every vertex in C is internal. Also, C does not contain any edge $e = \{x, y\}$ from G ; otherwise, there would be a cut F' that differentiates e , contradicting the assumption for case 2. Since every edge in C is in G_Δ , it must be the case that for every edge e in C there exists a cut F where $e \subseteq F_\cup$. Also, at most two vertices of C are in F_\cup . Thus the cuts $F^{(1)}, F^{(2)}$ and $F^{(3)}$ are distinct.

To simplify notation, for each $i \in \{1, 2, 3\}$ let $A_i = A_{F^{(i)}}$ and $B_i = B_{F^{(i)}}$. Without loss of generality, assume that $E(A_1) \cap F^{(2)} \neq \emptyset$ and $E(B_2) \cap F^{(1)} \neq \emptyset$. There are three possibilities.

- Suppose $F^{(3)} \cap E(A_2) \neq \emptyset$. If $x_1 \in A_2$, then by Lemma 18, $x_1 \in F_\cap^{(2)}$ and C is not chordless, a contradiction. Thus, $x_1 \in B_2$. Similarly, if $x_4 \in B_2$, by Lemma 18, $x_4 \in F_\cap^{(2)}$ and C is not chordless, a contradiction. Thus, $x_4 \in A_2$. Since C is a cycle, $F^{(2)}$ is a minimal cut and $\{F_\cup^{(2)} \setminus \{x_2, x_3\}\} \cap V(C) = \emptyset$, there exists an edge $\{v_1, v_2\}$ in C where $v_1 \in V(A_2) \setminus F_\cup^{(2)}$ and $v_2 \in V(B_2) \setminus F_\cup^{(2)}$. But, by Lemma 19, such an edge cannot exist.
- Suppose $F^{(3)} \cap E(A_1) \neq \emptyset$ and $F^{(3)} \cap E(B_2) \neq \emptyset$. Without loss of generality, assume that A_3, B_3 contain $F^{(2)}$ and $F^{(1)}$ respectively. Assume that $x_2 \in A_3$. Since $x_2 \in F_\cup^{(1)}$, by Lemma 18, $x_2 \in F_\cap^{(3)}$. Then, there exists an edge $\{x_2, x_4\}$ and C is not chordless, a contradiction. Thus, $x_2 \in B_3$. But $x_2 \in F_\cup^{(2)}$ and thus, by Lemma 18, $x_2 \in F_\cap^{(3)}$. Hence, there exists a chord $\{x_2, x_4\}$ and C is not chordless, again a contradiction.
- Suppose $F^{(3)} \cap E(B_1) \neq \emptyset$. Renaming vertices x_1, x_2, x_3 and x_4 as, x_4, x_3, x_2 and x_1 , respectively, brings us back to subcase 2(b).

Thus, G_Δ does not contain a chordless cycle of length four or greater; hence, G_Δ is chordal. \square

Proof of Theorem 11. Lemma 21 states that G_Δ is triangulated. We now prove that G_Δ is a legal triangulation; i.e., that it satisfies conditions (LT1) and (LT2) of Section ‘Display graphs and edge label intersection graphs’.

Condition (LT2) holds for G_Δ , because our construction adds no fill-in edge incident on a leaf. Now suppose that G_Δ violates (LT1); i.e., G_Δ has a clique H with two internal edges $e = \{x_1, y_1\}$ and $e' = \{x_2, y_2\}$. Let F be the cut that differentiates e . Assume that $x_1 \in V(A_F)$ and $y_1 \in V(B_F)$. By Lemma 20, F_\cup does not contain both endpoints of e' . Without loss of generality, assume that $x_2 \notin F_\cup$ and $x_2 \in$

A . Since $x_2 \notin F_\cup$ and $y_1 \notin F_\cap$, by Lemma 19, there is no edge between x_2 and y_1 in G_Δ . Thus, H is not a clique of G_Δ , a contradiction. Hence, G_Δ satisfies (LT1) and is therefore a legal triangulation of $G(\mathcal{P})$. \square

Conclusion

We have shown that the characterization of tree compatibility in terms of restricted triangulations of the edge label intersection graph transforms into a characterization in terms of minimal cuts in the display graph. These two characterizations are closely related to the legal triangulation characterization of [8]. We also derived characterizations of the agreement supertree problem in terms of minimal cuts and minimal separators of the display and edge label intersection graphs respectively.

It remains to be seen whether any of our characterizations can lead to explicit fixed-parameter algorithms for the tree compatibility and agreement supertree problems when parametrized by the number of trees. Indeed, as of yet, the fixed-parameter tractability of agreement remains open.

We close with some remarks on characterizations of two problems related to compatibility. A profile \mathcal{P} defines a tree S if S is the only compatible supertree for \mathcal{P} . \mathcal{P} identifies a tree S if S is a compatible supertree for \mathcal{P} and every other compatible supertree for \mathcal{P} displays S . Grunewald et al. [16] use quartet graphs to characterize when a profile consisting of quartet trees defines or identifies a tree. An interesting question is whether similar characterizations can be derived for arbitrary profiles using display graphs or edge label intersection graphs. Along these lines, we note a connection between complete sets of cuts and the question of whether a profile defines a tree, which was pointed out by one of the reviewers. To explain it, we need some definitions ([10], p. 131). Let T be a tree and let $q = xy|wz$ be a quartet tree displayed by T . Quartet tree q distinguishes an interior edge e of T if e is the only interior edge such that $\{x, y\}$ and $\{w, z\}$ are in different connected components of $T - e$. Now, let S and T be two trees such that S displays T . An interior edge e of T distinguishes an interior edge f of S if there exists a quartet q such that e and f are both distinguished by q . Suppose \mathcal{P} is a profile in which there is at least one taxon in common among all input trees. Then, \mathcal{P} defines a tree S if and only if \mathcal{P} is compatible and every interior edge of S is distinguished by an interior edge of at least one tree in \mathcal{P} ([10], p. 133). Now, recall that if \mathcal{F} is a complete set of cuts of $G(\mathcal{P})$, then, for every tree $T_i \in \mathcal{P}$ and every internal edge e of T_i , there is some cut $F \in \mathcal{F}$ in which e is the only edge of T_i . Thus, if \mathcal{P} is compatible, e must be a distinguishing edge for some internal edge of a supertree for \mathcal{P} . This observation could lead to a cut-based characterization of definability analogous to known triangulation-based characterizations (see [10], p. 79).

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

SV stated and proved the main results of the paper and wrote most of the first draft. DFB proposed the research topic to SV, supervised the research, contributed to the first draft, and was in charge of the final draft. Both authors read and approved the final manuscript.

Acknowledgements

We thank Sylvain Guillemot for his valuable comments. We are also grateful to the reviewers for providing constructive criticism. This work was supported in part by the National Science Foundation under grants CCF-1017189 and DEB-0829674.

Received: 2 December 2013 Accepted: 11 April 2014

Published: 17 April 2014

References

1. Gordon AD: **Consensus supertrees: the synthesis of rooted trees containing overlapping sets of labelled leaves.** *J Classif* 1986, **9**:335–348.
2. Aho A, Sagiv Y, Szymanski T, Ullman J: **Inferring a tree from lowest common ancestors with an application to the optimization of relational expressions.** *SIAM J Comput* 1981, **10**(3):405–421.
3. Ng M, Wormald N: **Reconstruction of rooted trees from subtrees.** *Discrete Appl Math* 1996, **69**(1–2):19–31.
4. Steel MA: **The complexity of reconstructing trees from qualitative characters and subtrees.** *J Classif* 1992, **9**:91–116.
5. Bryant D, Lagergren J: **Compatibility of unrooted phylogenetic trees is FPT.** *Theor Comput Sci* 2006, **351**:296–302.
6. Courcelle B: **The monadic second-order logic of graphs I, Recognizable sets of finite graphs.** *Inf Comput* 1990, **85**:12–75.
7. Arnborg S, Lagergren J, Seese D: **Easy problems for tree-decomposable graphs.** *J Algorithms* 1991, **12**(2):308–340.
8. Vakati S, Fernández-Baca D: **Graph triangulations and the compatibility of unrooted phylogenetic trees.** *Appl Math Lett* 2011, **24**(5):719–723.
9. Gysel R, Stevens K, Gusfield D: **Reducing problems in unrooted tree compatibility to restricted triangulations of intersection graphs.** In *Algorithms in Bioinformatics – 12th International Workshop, WABI 2012 Ljubljana, Slovenia, September 10–12, 2012. Proceedings, Volume 7534 of Lecture Notes in Computer Science.* Edited by Raphael BJ, Tang J. Heidelberg: Springer; 2012:93–105.
10. Semple C, Steel M: *Phylogenetics.* Oxford Lecture Series in Mathematics, Oxford: Oxford University Press; 2003.
11. Buneman P: **The recovery of trees from measures of dissimilarity.** In *Mathematics in the Archaeological and Historical Sciences.* Edinburgh: Edinburgh University Press; 1971:387–395.
12. Parra A, Scheffler P: **Characterizations and algorithmic applications of chordal graph embeddings.** *Discrete Appl Math* 1997, **79**(1–3):171–188.
13. Bouchitté V, Todinca I: **Treewidth and minimum fill-in: grouping the minimal separators.** *SIAM J Comput* 2001, **31**:212–232.
14. Heggernes P: **Minimal triangulations of graphs: a survey.** *Discrete Math* 2006, **306**(3):297–317.
15. Gusfield D: **The multi-state perfect phylogeny problem with missing and removable data: solutions via integer-programming and chordal graph theory.** *J Comput Biol* 2010, **17**(3):383–399.
16. Grunewald S, Humphries PJ, Semple C: **Quartet compatibility and the quartet graph.** *Electron J Comb* 2008, **15**:R103.

doi:10.1186/1748-7188-9-13

Cite this article as: Vakati and Fernández-Baca: Characterizing compatibility and agreement of unrooted trees via cuts in graphs. *Algorithms for Molecular Biology* 2014 **9**:13.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

